

**Changement linguistique et périodisation du français
(pré)classique : deux études de cas à partir des corpus du
RCFC**

Journal:	<i>Journal of French Language Studies</i>
Manuscript ID	JFL-AR-2019-0018.R4
Manuscript Type:	Article
Keywords:	Changement linguistique, périodisation, français (pré)classique, recatégorisation des formes dedans / dessous / dessus / dehors, montée des clitiques, Réseau Corpus Français Préclassique et Classique (RCFC)

SCHOLARONE™
Manuscripts

Changement linguistique et périodisation du français (pré)classique : deux études de cas à partir des corpus du RCFC

Résumé

Cette étude propose d'explorer et d'identifier des moments particuliers où le changement linguistique se produit, afin de confirmer ou de rejeter l'idée d'une période spécifique désignée par le terme « français préclassique », avec une rupture – ou frontière chronolectale – détectable autour de 1630 (cf. Ayres-Bennett et Caron, 2016). Afin de vérifier dans quelle mesure cette chronologie peut être confirmée, il est nécessaire de multiplier des analyses fines et pointues sur des traits linguistiques qui ont subi des changements à l'époque en question et d'interroger une gamme de textes qui reflètent la variation discursive et pragmatique, au lieu de consulter le canon des traditions textuelles actuellement disponibles sur des bases numérisées, qui sont essentiellement littéraires. C'est pourquoi nous avons consulté des sources de nature différente, qui pourraient attester des usages émergents, à savoir les corpus du Réseau Corpus Français Préclassique et Classique (RCFC). Seront présentés les résultats de deux études de cas (la recatégorisation des formes *dedans* / *dessous* / *dessus* / *dehors* et la montée des clitiques), abondamment discutés par les remarqueurs.

1. Introduction

1.1 Problématique de la périodisation du français (pré-classique) : état de la question

À l'intérieur du modèle tripartite traditionnel de périodisation de la langue française (ancien français, moyen français, français moderne) des subdivisions ont été introduites, qui ont amené à distinguer un français classique et un français préclassique. Cependant le consensus sur la pertinence de différencier cette dernière période n'est pas général. C'est notamment la définition des limites chronologiques du français préclassique et, en conséquence, l'émergence du français classique, qui est débattue.¹

Nous proposons d'explorer et d'identifier des moments particuliers où le changement linguistique se produit, pour confirmer – ou rejeter – l'idée d'une période spécifique désignée par le terme « français préclassique ». Selon des investigations récentes (Ayres-Bennett et Caron, 2016), une rupture – ou frontière chronolectale, pour utiliser le terme proposé par Caron (2002) – serait détectable autour de 1630.

Afin de vérifier dans quelle mesure cette chronologie peut être confirmée, il est nécessaire, comme cela a été recommandé par Combettes et Marchello-Nizia (2010), de multiplier des analyses fines et pointues sur des traits linguistiques qui ont subi des changements à l'époque en question. Le présent travail, dont l'objectif est de discuter les résultats de deux études de cas (la recatégorisation des formes *dedans* / *dessous* / *dessus* / *dehors* et la montée des clitiques), va dans cette direction, en contribuant à enrichir nos connaissances sur les étapes et les orientations de l'évolution du français. Il s'agit d'étudier en détail deux « micro-changements » ('micro-changes' en anglais) de la grammaire, selon la terminologie de Traugott et Trousdale (2013), qui constituent des cas d'une fréquence assez importante pour permettre une interrogation sur corpus.

Étant donné qu'en général la périodisation est déterminée par l'évolution et la fixation de la norme (Blake, 1994 : 42), il est légitime de se demander si la même périodisation peut être obtenue sur des bases de textes de nature différente de ceux habituellement consultés. Plus précisément, il nous a paru souhaitable d'interroger une gamme de textes qui reflètent la variation discursive et pragmatique, en élargissant le canon des traditions textuelles actuellement disponibles sur des

¹ Pour un état complet de la question, cf. Ayres-Bennett et Caron (2016).

bases numérisées, qui sont essentiellement littéraires. Cette étude servira ainsi de complément aux résultats déjà obtenus pour l'usage littéraire, par exemple, par Yvette Galet (1971) dans son excellente étude sur la montée des clitiques (cf. section 3). En faisant l'analyse des sources de nature différente, qui pourraient attester des usages émergents, nous avons la possibilité de voir si on arrive à une périodisation différente, et notamment d'examiner si l'usage littéraire est plutôt conservateur par rapport aux genres plus informels, plus spontanés, ou plus proches de l'oral. Les corpus du Réseau Corpus Français Préclassique et Classique (RCFC), que nous avons retenus pour nos recherches, satisfont ces conditions et répondent ainsi à nos exigences. Le RCFC, créé en 2014, met en commun cinq différents projets d'étude de diachronie du français (SERMO, QUINTE-CURCE, BAILLY, APWCF, ULFELDT) que nous décrivons brièvement ci-dessous.²

1.2 Description des corpus

Le corpus SERMO est composé de sermons protestants, édités à Genève pour la plupart, entre 1550 et 1750. Il comprend 62 sermons et environ 600 571 mots. Le corpus QUINTE-CURCE, consiste en cinq traductions françaises des *De regibus gestis Alexandri Magni regis Macedonum*, de Quinte-Curce, une par Nicolas Séguier datant de 1598, une par Nicolas de Soulfour de 1629, une par Bernard Lesfargues de 1639,³ et les deux dernières par Claude Favre de Vaugelas, qui y travailla pendant trente ans, jusqu'à sa mort, en la corrigeant et la remaniant constamment, et dont deux versions ont été publiées, posthumes, l'une en 1653 et l'autre en 1659. Les livres trois et quatre du texte, qui font l'objet des analyses, atteignent, dans les quatre versions retenues, au total 145 525 mots. Le corpus BAILLY est constitué de la correspondance qu'Albert Bailly (1605-1691), barnabite savoyard, agent officieux de la duchesse de Savoie à Paris pendant la Fronde et évêque d'Aoste ensuite, adressa à la cour de Turin entre 1643 et 1688. Il réunit 945 lettres et gazettes (environ 601 000 mots) envoyées de France (essentiellement de Paris), de Savoie et d'Aoste entre 1643 et 1688. Le corpus APWCF, *Acta Pacis Westphalicae*, contient une grande partie de la correspondance française échangée entre les représentants de la France à Münster et le

² Les projets sont présentés de manière plus approfondies dans Amatuzzi *et al.* (2019).

³ Ce texte se distingue nettement des autres traductions par sa longueur et la façon de traduire adoptée et il ne fait pas donc partie de l'analyse ici.

gouvernement français à Paris pendant les négociations concernant la Paix de Westphalie (1648). Il compte 2,4 millions de mots dans 1670 lettres écrites entre 1644 et 1647, pour la plupart à Münster (738) et à Paris (506). Enfin, le corpus ULFELDT prend la forme d'une lettre privée (24 000 mots) contenant une autoprésentation rédigée en 1673 par la princesse danoise Leonora Christina (1621–1698), fille du roi danois Christian IV, lorsqu'elle se trouvait en prison. Il s'agit d'une autodéfense en vue de sa libération, car le destinataire Otto Sperling (junior) avait prévu de se servir du texte pour intervenir en faveur de la princesse auprès des Cours européennes. Sauf dans le cas de QUINTE-CURCE, qui sert de point de référence de la norme, les documents de tous les corpus ont été rédigés dans des situations de communication ancrées dans des contextes concrets et réels et appartiennent à des genres textuels variés, ce qui nous permet de retracer la diffusion de changements dans des types de texte différents.

Nous nous intéressons à cette gamme de textes comprenant à la fois des « ego-documents »,⁴ considérés comme des sources proches de l'oral, des sermons et des correspondances qui utilisent parfois une variété linguistique plus formelle, ainsi qu'un corpus plus littéraire (traduction et narration historique). Chaque corpus n'est évidemment pas homogène,⁵ mais si l'on les catégorise dans les grandes lignes selon cinq des critères proposés par Koch et Oesterreicher (2001 : 586, première version 1985) pour distinguer les traits caractéristiques de la communication « proche » ou « distante » – (a) communication privée / publique ; (b) ancrage / détachement actionnel et situationnel ; (c) co-présence / séparation spatio-temporelle ; (d) coopération communicative intense/minime ; (e) communicative spontanée / préparée – les diversités sont évidentes (cf. Tableau 1) :

[INSÉRER LE TABLEAU 1 ICI]

⁴ Ce terme est employé pour indiquer les écrits dans lesquels le « je », l'auteur, est toujours présent dans le texte comme sujet parlant. Les « ego-documents » comprennent les lettres ou des ouvrages autobiographiques tels que les mémoires, les journaux et les récits de voyage. Pour citer Nevalainen et Raumolin-Brunberg (2012 : 32), les lettres représentent ce qui se rapproche le plus de l'oral authentique.

⁵ Pour une discussion critique du problème de la non-homogénéité d'un corpus, cf. Gerstenberg (2018).

	Communication		Actionnel et situationnel		Spatio-temporel		Coopération communicative		Communication	
	privée	publique	ancrage	détachement	co-présence	séparation	intense	minime	spontanée	préparée
SERMO		+	+		+	+		+		+
QUINTE-CURCE		+		+		+		+		+
BAILLY	+	+	+			+	+	+	+	
APWCF	+	+	+	+		+		+	+	+
ULFELDT	+		+			+	+	+	+	

Tableau 1 : Sélection des paramètres de la communication (Koch et Oesterreicher, 2001 : 586)

Nous avons ainsi d’une part le corpus QUINTE-CURCE, qui contient des traductions exclusivement écrites, rédigées dans un milieu plutôt érudit, sans échange direct ou personnel avec un public visé, et qui représente nettement la distance. L’intérêt linguistique de ces traductions réside dans le fait qu’à la période où elles ont été réalisées, la traduction avait un statut de création littéraire et était regardée comme une sorte de « grammaire appliquée », c’est-à-dire une réalisation pratique de la grammaire ou de la norme (cf. Ayres-Bennett et Caron, 1996). ULFELDT se trouve d’autre part proche du pôle de la proximité. En tant que lettre privée contenant une autoprésentation, le texte du corpus ULFELDT ne semble pas suivre un modèle générique particulier. Il s’agit d’un récit chronologique de la vie de Leonora Christina, mariée Ulfeldt. Au niveau explicite, le rôle de la princesse est de fournir des informations apparemment anodines ; en réalité, il s’agit d’un texte hautement politique destiné à la diffusion par l’allocutaire. Le texte incarne certes une communication à distance, parce qu’écrit en l’absence du destinataire, mais, en tant qu’échange strictement confidentiel, de nature privée et amicale, il contient un grand nombre de traits caractéristiques de l’immédiat communicatif et beaucoup de traces linguistiques d’oralité. Malgré la présence d’interférences linguistiques avec le danois et d’erreurs de grammaire et tout en ayant été rédigé en situation d’extraterritorialité, le texte révèle une bonne connaissance de la langue française de la part de la princesse.

Les trois autres corpus possèdent un mélange de traits de distance et de proximité. Les deux corpus de lettres (BAILLY, APWCF) se distinguent. Les textes du corpus BAILLY sont des lettres et des gazettes écrites pour relater des informations obtenues, demander des faveurs, donner des conseils. Puisque les destinataires sont toujours des personnes connues mais physiquement éloignées, les lettres attestent d’une communication à distance entre deux

personnes de confiance. Bailly se soumet au code et au cérémonial ritualisés de la communication épistolaire, contraignants au plan sociolinguistique. Avec le corpus APWCF c'est plutôt le genre *lettre diplomatique* qui est représenté. Dans cette correspondance, les diplomates décrivent l'avancement des négociations, en cherchant à convaincre les membres du gouvernement à Paris d'accepter les renseignements qu'ils leur fournissent. De façon plus directe, les lettres envoyées depuis Paris donnent des instructions aux diplomates, en les appelant parfois à l'ordre. Quant au corpus SERMO, la finalité des sermons est d'« améliorer la compréhension de la doctrine et d'inciter les fidèles à mener une vie plus conforme aux exigences religieuses » (Maingueneau, 2009: 39), alternant de fait entre enseignement et argumentation. Si les fidèles restent muets, le prédicateur tient néanmoins compte de leurs compétences et adapte son discours au peuple. Le rapport des sermons avec la langue orale est très étroit mais, comme le sermon peut avoir été rédigé avant ou après avoir été prononcé, de manière très fidèle ou complètement réécrite, certains textes se caractérisent par des traits typiques de l'oral spontané, alors que d'autres sont fort élaborés, suivant un plan de composition très scolaire, parfois même noté dans la marge.

1.3 Les potentialités des corpus RCFC

Quelles sont les potentialités des corpus du RCFC, qui ont par ailleurs l'avantage d'être interrogeables à travers des ressources informatiques⁶ appropriées dont les modèles informatiques ont été adaptés aux exigences de la langue du 17^e siècle ? Compte tenu de la pluralité des genres représentés, ces corpus donnent l'opportunité d'observer la diffusion de traits linguistiques dans un univers textuel riche et sociologiquement diversifié. Le recensement des traits et des variables linguistiques caractéristiques nous fournira des renseignements précieux sur l'élaboration et l'établissement de la norme. Pour ce qui est plus particulièrement du problème de la périodisation, il est possible d'évaluer d'une part si les changements linguistiques se font à des rythmes semblables ou inégaux dans les genres disponibles, et, d'autre part, si les spécificités génériques et pragmatiques peuvent expliquer les variations dans l'évolution du français.

⁶ On trouvera les références des sites et des moteurs de recherche utilisés pour l'étude dans la liste des références à la fin de l'article.

Dans cette optique nous avons choisi de nous concentrer sur deux cas de variation suivie d'un changement, parmi d'autres réputés saillants puisqu'ils ont été discutés par les remarqueurs, c'est-à-dire les auteurs des volumes d'observations sur le « bon usage » du français (Ayres-Bennett et Seijido, 2011). Ceux-ci ont donné leur avis sur leur usage, ce qui montre que ces changements en cours étaient ressentis comme des cas de variation problématiques, dont l'emploi devait être réglé. Le premier (section 2) est la *recatégorisation* des formes plurifonctionnelles (adverbes et prépositions) *dedans / dessous / dessus / dehors*, qui ne conservent vers la fin de la période qu'une fonction unique, la fonction adverbiale. Ce cas n'est pas à proprement parler un cas de grammaticalisation dans le sens traditionnel du terme, à savoir une réanalyse d'un lexème qui va obtenir un statut grammatical (Lehmann, 2015). Il s'agit plutôt d'un cas de « regrammation » selon la terminologie d'Andersen (2008), c'est-à-dire la réanalyse d'unités déjà grammaticales, le cas échéant plurifonctionnelles, qui sont réanalysées en tant que formes exclusivement adverbiales. Le deuxième cas (section 3) concerne la montée des clitiques. C'est un changement au niveau de la construction, qui résulte en une réduction de la cohésion interne entre un verbe semi-auxiliaire et son infinitif, réduction qui se manifeste par le changement de position du clitique, comme *je le veux voir*, qui est progressivement remplacé par la structure *je veux le voir*. Ce changement pourrait être considéré comme une dégrammaticalisation, dans la mesure où le verbe régissant perd son statut de (semi-)auxiliaire, ou bien, puisqu'il s'agit d'une perte de la nature grammaticale d'une construction, un cas de « déconstructionnalisation ». Une autre raison pour choisir ces deux cas réside dans le fait que le premier semble faire partie de toute une série de changements qui, après une période d'évolution rapide, voit l'élimination complète ou presque complète de l'ancienne variante avant 1630, tandis que le deuxième représente plutôt un exemple d'une évolution qui commence à s'accélérer autour de 1630 (Ayres-Bennett et Caron, 2016 : 44-47). Ils représentent ainsi deux aspects différents de la frontière chronolectale à l'étude.

En traitant ces deux cas, nous considérons l'importance des facteurs temps, espace, genre textuel et type de communication pour la diffusion des changements. Nous défendons l'idée que la diffusion du changement étudié se poursuit non seulement dans le temps, mais aussi dans l'espace selon un mouvement centrifuge, à partir du centre d'innovation que constitue Paris. Il est

donc possible que les textes rédigés loin du centre (ULFELDT, BAILLY, SERMO) soient en retard par rapport aux textes rédigés près du centre (QUINTE-CURCE, APWCF). D'autre part, il est prévisible que les textes non-littéraires soient en avance par rapport aux textes littéraires.

2. Première étude de cas : recatégorisation des unités plurifonctionnelles, adverbess et prépositions

2.1 Introduction

Entre la fin du 16^e siècle et la première moitié du 17^e siècle, le français procède à une réorganisation des parties du discours (Fournier, 1998 : 139-141 ; Combettes et Marchello-Nizia, 2010) selon deux tendances : de nouvelles catégories grammaticales sont créées et des unités ayant plusieurs fonctions n'en conservent plus qu'une seule. Certains cas sont bien connus et étudiés : les possessifs et les démonstratifs, par exemple, dont les formes, à l'origine plurifonctionnelles, se réorganisent autour des deux catégories parallèles de pronoms et déterminants, se spécialisant dans l'une ou l'autre fonction (Marchello-Nizia, 1995 : chapitre 5 ; Fournier, 1998 : 139-141).

Nous examinons ici la distinction progressive entre adverbess et prépositions et le processus de recatégorisation des formes *dedans* / *dessous* / *dessus* / *dehors*. Jusqu'au 16^e siècle ces formes pouvaient être utilisées indistinctement comme prépositions et, en emploi absolu, comme adverbess (Marchello-Nizia, 1999 : 120 ; 2002). Elles se superposaient donc aux formes simples *dans* / *sous* / *sur* / *hors* ; en particulier *dans* est rare avant le 17^e siècle et *dedans* plus répandu (Fagard et Combettes, 2013 ; Vigier, 2015). Au cours du 17^e siècle une modification de l'inventaire catégoriel du français s'opère avec le résultat que les formes composées se spécialisent en tant que adverbess, alors que les formes simples ne sont utilisées que comme prépositions. C'est ainsi que *dans* remplace rapidement *dedans* (Fagard et Sarda, 2009 : 222 ; Vigier, 2017: 107, 117). Le résultat de cette évolution est conforme aux recommandations formulées dans la remarque consacrée à *sur* / *sous* par Vaugelas dans ses *Remarques sur la langue française* (1647), dont le but est de fixer l'usage encore fluctuant. Il observe :

Ces prepositions se doivent tousjours mettre simples, si ce n'est en certains cas que nous remarquerons. Je les appelle simples en comparaison des composées *dessus*, & *dessous*, que tout le monde presque employe indifferemment, & en prose, & en vers, pour *sur*, & *sous*. On en fait autant de quelques autres prepositions, comme *dedans*, *dehors* [...]. Il semble que ces composez soient plustost adverbess que prepositions ; car leur grand usage est à la fin des periodes, sans rien regir après eux, puis qu'ils terminent la periode & le sens [...]. Au lieu que les prepositions sont perpetuellement suivies d'un nom, ou d'un verbe, ou de quelque autre partie de l'Oraison, comme le porte le nom mesme de preposition. (Vaugelas, 1647: 124–125)⁷

Vaugelas précise ensuite qu'il y a trois cas exceptionnels où la préposition simple doit être remplacée par la préposition composée (Vaugelas, 1647 : 125-126) :

- quand on met les deux contraires ensemble, & tout de suite, comme *Il n'y a pas assez d'or ni dessus, ni dessous la terre* [...]
- quand il y a deux prepositions de suite, encore qu'elles ne soient pas contraires, comme *elle n'est ni dedans, ni dessus le coffre*
- lors qu'il y a une autre préposition devant, comme *il luy a paßé par dessus la teste*.

L'objectif de l'interrogation de nos corpus est de mieux situer chronologiquement ce cas de *regrammation* qui représente la dernière étape d'un long processus de transformation qui a touché le système prépositionnel et qui aurait comporté une réanalyse des propriétés syntaxiques de plusieurs prépositions complexes, lesquelles ont cessé de pouvoir régir un complément (cf. Tremblay *et al.*, 2003 : 560-561).

2.2 Exploration du corpus, remarques préliminaires

Au moment d'interroger les corpus pour obtenir toutes les occurrences des formes *dans* / *dedans*, *sur* / *dessus*, *sous* / *dessous* et *hors* / *dehors* nous avons été confrontées à un problème concernant l'étiquetage POS (Part-of-speech tagging), car les outils informatiques dont nous disposons ont été

⁷ Nous respectons fidèlement l'orthographe originale de nos textes, mais nous avons modernisé l'usage de *u* / *v* et de *i* / *j*.

développés dans le cadre d'autres projets et doivent être adaptés aux besoins spécifiques de l'état de langue qui nous intéresse. Dans le cas précis, le Tree-tagger que nous utilisons, élaboré par PRESTO⁸ à partir d'un lexique fourni par LGeRM (ATILF),⁹ n'est pas en mesure de distinguer les différentes fonctions des formes plurifonctionnelles *dedans*, *dessus*, *dessous*, *dehors* en leur attribuant des POS distinctes. De ce fait, nous avons relevé toutes les occurrences de ces formes et nous avons ensuite désambiguïsé les résultats manuellement. Cela n'a évidemment pas été nécessaire pour les formes *dans* / *sur* / *sous*. Or, la taille différente des corpus s'est révélée un point de force pour l'implémentation du questionnement scientifique car nous avons pu travailler sur deux plans complémentaires : nous avons effectué des enquêtes à la main et des vérifications complètes sur les corpus plus petits et, en conséquence, nous avons pu adapter et préciser les investigations automatiques lancées sur les corpus plus vastes (cf. AmatuZZi *et al.*, 2019).

2.3 Résultats de l'exploration du corpus

2.3.1 Identification des facteurs en jeu

Nous avons recensé tous les cas où, dans nos corpus, les formes complexes composées *dedans* / *dessous* / *dessus* / *dehors* étaient employées avec valeur prépositionnelle, en les mettant en relation avec l'emploi des formes simples correspondantes *dans* / *sous* / *sur* / *hors*. Ces données quantitatives doivent rendre possible une évaluation de la fréquence de ces formes dans le temps, en vue de déterminer s'il y a un moment précis où les formes composées sont soumises à une désaffectation en tant que prépositions et si cela se fait au profit des formes simples.¹⁰

Pour harmoniser la période explorée, nous nous sommes limitées aux textes publiés avant 1673. Pour le corpus SERMO, qui présente une diachronie longue, trois tranches temporelles (1590-1630, 1631-1642, 1643-1673) ont été distinguées.¹¹ Le choix chronologique a été déterminé dans le but d'une part de mieux considérer la date de 1630, proposée par Ayres-Bennett et Caron

⁸ Projet franco-allemand (<http://presto.ens-lyon.fr>).

⁹ LGeRM (Lemmes Graphiques et Règles Morphologiques): lemmatiseur conçu pour gérer la variation graphique historique du français. Développé pour le moyen français, adapté au français du 16^e / 17^e siècle. <http://www.atilf.fr/LGeRM>).

¹⁰ Une chronologie quelque peu différente est offerte par Odile Leclercq (2018).

¹¹ Les citations de la Bible ont été exclues des calculs parce qu'elles représentent un langage archaïque (Skupien-Dekens, 2018).

(2016) comme moment de démarcation entre français préclassique et français classique, et d'autre part d'isoler la période 1643-1673, ce qui permet une comparaison plus pertinente avec les autres corpus, et notamment avec celui de BAILLY.

Les résultats obtenus ont été triés selon les critères suivants :

- ont été omis tous les cas de nominalisation, comme les exemples suivants: *Les Espagnols reprennent le dessus* (APWCF) ; *Tout le dessus estoit de Votre Altesse Royale* (BAILLY); *Dieu luy mesme paroissoit au dessus* (SERMO).
- ont été exclues les occurrences des formes étudiées lorsqu'elles sont précédées de prépositions ou répétées, tel *exalté par dessus tous les cieux* (SERMO) et *Il a des intelligences dedans et dehors le pays* (BAILLY), car cette utilisation, comme l'avait ressenti Vaugelas (cf. supra), est différente. Il s'agit de constructions prépositionnelles et / ou adverbiales qui ont subi un processus de grammaticalisation complexe et spécifique à cause de la nature même des noms qu'elles comportent (De Mulder, 2003 ; Diedihou, 2011 ; Fagard et De Mulder, 2007). De plus, ces formes ne subissent pas de variation dans la période qui nous intéresse.
- ont été spécifiés les cas de *hors* et de *hors de* par souci de précision, même si Vaugelas (1647 : 124) affirme que *tous deux sont bons*.¹²

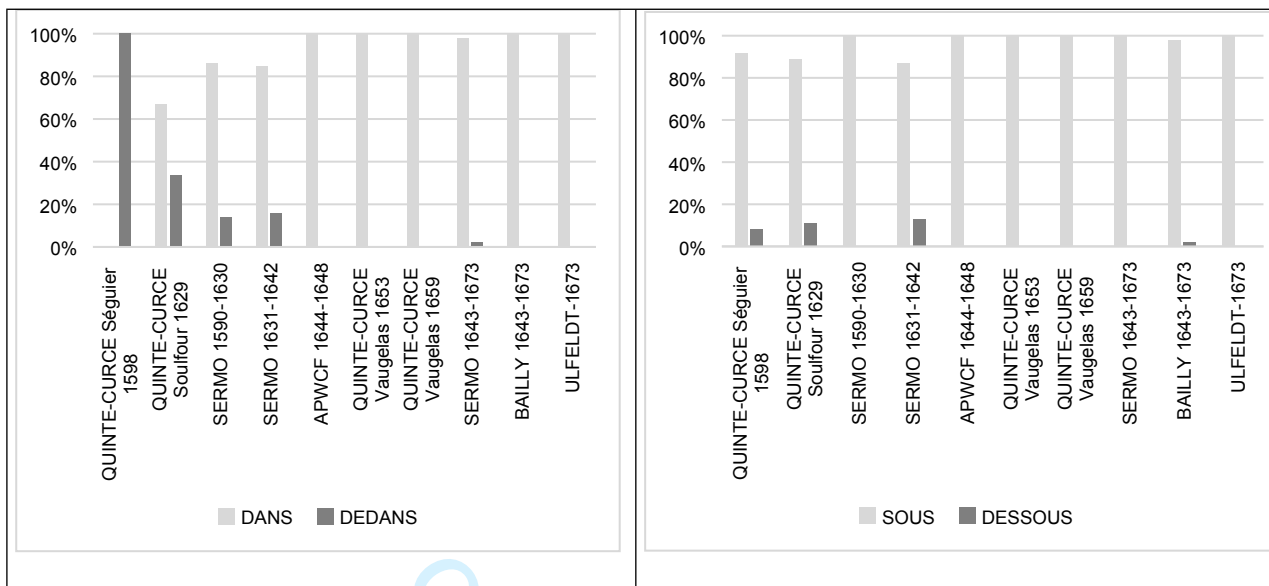
2.3.2 Statistiques

Nous présentons dans le Tableau 2 et la Figure 1 les résultats de l'interrogation sur les sous-corpus.

[INSERER LE TABLEAU 2 et LA FIGURE 1 ICI]

¹² Nous avons également fait le calcul des usages adverbiaux, mais les résultats se sont révélés sans grand intérêt pour notre étude.

Tableau 2 : Formes en fonction de préposition



	QUINTE-CURCE Séguier 1598	QUINTE-CURCE Souffour 1629	SERMO 1590-1630	SERMO 1631-1642	APWCF 1644-1648	QUINTE-CURCE Vaugelas 1653	SERMO 1643-1673	BAILLY 1643-1673	QUINTE-CURCE Vaugelas 1659	ULFELDT 1673
<i>Dans + N ou Pro</i>	0	100	31	60	11677	176	527	1705	174	75
<i>Dedans + N ou Pro</i>	64	50	5	11	6	0	12	0	0	0
<i>Sous + N ou Pro</i>	11	8	33	20	216	30	82	51	21	2
<i>Dessous + N ou Pro</i>	1	1	0	3	0	0	0	1	0	0
<i>Sur + N ou Pro</i>	109	149	140	43	8421	162	336	1066	143	71
<i>Dessus + N ou Pro</i>	0	17	1	6	9	0	0	0	0	0
<i>Hors + N ou Pro</i>	8	6	3	1	70	2	2	4	2	2
<i>Hors de + N ou Pro</i>	1	6	9	4	303	9	12	49	10	8
<i>Dehors + N ou Pro</i>	0	0	0	0	1	0	0	0	0	0

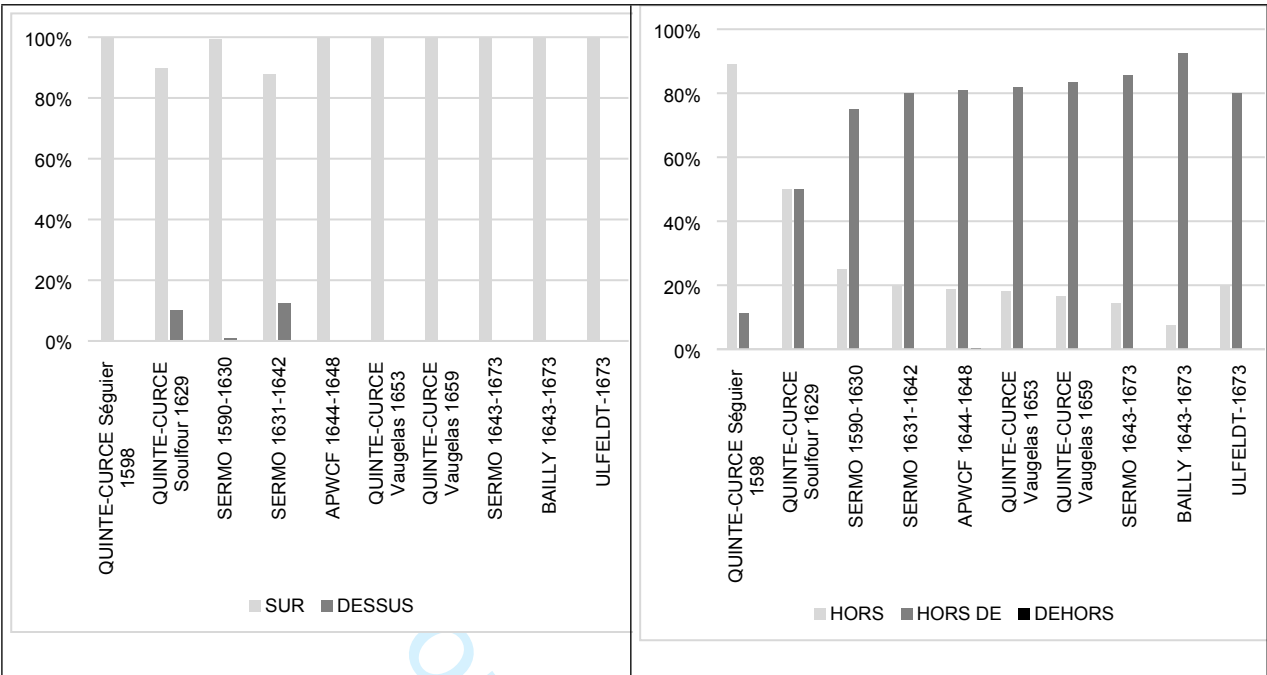


Figure 1 : Formes en fonction de préposition¹³

2.4. Analyses

Même si les chiffres bruts révèlent une quantité limitée d’attestations dans certains corpus (Tableau 2), l’analyse de toutes les données présentées dans la Figure 1 permet d’avoir des informations intéressantes du point de vue qualitatif. On observe tout d’abord que les formes composées *dedans*, *dessous*, *dessus* sont de moins en moins utilisées en fonctions prépositionnelle. Cette baisse est compensée par la croissance de l’emploi des formes simples en fonction de préposition.

Le changement le plus évident concerne *dedans*. Si nous laissons de côté les statistiques pour SERMO, le niveau d’attestation de l’ancien usage chute nettement après Souffour 1629 (33,3%) et devient pratiquement nul dans tous les corpus avec seulement une poignée d’attestations dans APWCF (0,1%). Les périodes SERMO 1590-1630 (13,9%) et SERMO 1643-1673 (2,2%) se conforment à notre hypothèse selon laquelle le changement est terminé autour de 1630, mais le pourcentage plutôt élevé de *dedans* comme préposition (15,5%) pour la tranche temporelle 1631-1642 paraît surprenant. Il peut être expliqué par la distribution des occurrences. En effet, 8 des 11 emplois prépositionnels de *dedans* se trouvent chez un seul prédicateur, Ezekiel

¹³ On trouve une seule occurrence de *dehors* comme préposition, dans le corpus APWCF 1644-1648. La fréquence est trop faible pour apparaître dans la figure.

Marmet, dont les sermons furent publiés en 1636. Il s'agit essentiellement d'expressions plus ou moins figées, comme par exemple : *des espines qui estouffent dedans les cœurs humains; tout l'espoir ils aiment mieux vivre esclaves en terre que Rois dedans le ciel; Ils l'ont transporté dedans les cieux, nostre ame est retenue dedans la prison de ce corps*, qui font partie de ce qu'on appelle « le patois de Canaan » (Skupien-Dekens, 2017 : 2018). Celles-ci reprennent le style des citations bibliques, qui « déteignent » sur le reste du texte. Pareillement, les trois autres occurrences se trouvent dans un sermon de Michel Le Faucheur (1632) qui emploie à deux reprises l'expression *dedans le Paradis* et une fois l'expression *dedans l'Evangile*.

Le même chemin de changement se répète pour *dessous*, qui est encore présent comme préposition dans 11,1% des cas dans Soulfour 1629 (*mais avec un sablon qui fuit dessous les pieds*), mais n'est plus enregistré ensuite, sauf pour 1,9 % dans BAILLY, avec une seule occurrence (*portant des armes dessous sa robe* 1672) et 3 exemples dans SERMO 1631-1642, dont deux sont de nouveau chez Ezekiel Marmet 1636 (*loger dessous le mesme toit ; recouvrer à chacun dessous son climat*).¹⁴ La tendance est la même pour *dessus* : après 10,2% d'occurrences dans Soulfour 1629 (*qu'ils empiètent quantité des vaisseaux qui viennent eschoüer dessus leurs sables*), les pourcentages restent inférieurs à 1%, à l'exception près de SERMO 1631-1642 où encore une fois c'est Ezekiel Marmet 1636 qui se sert de l'ancienne construction (*Elles attirent plusieurs maux dessus l'ame ; un commandement absolu dessus ses affections*).

Le cas de *dehors* se distingue légèrement des autres, car il ne figure pratiquement jamais en tant que préposition (seul le corpus APWCF enregistre une fois *dehors la ville*) alors que la forme simple *hors* est très répandue, et paraît de temps à autre comme élément des locutions verbales.¹⁵ Nous observons pourtant que *hors* + N ou Pro reste très fréquent jusqu'à Soulfour 1629 (*Car il estoit capable de souffrir tout hors les retardemens*), après lequel il devient une variante minoritaire, et même très minoritaire chez BAILLY (*hors la ville* 1651 ; *il n'est personne hors la vôtre* 1666 ; *hors le service* 1668 ; *hors les crimes* 1669).

¹⁴ Selon Ayres-Bennett et Caron, 2016, *dedans* est plus lent à disparaître comme préposition que *dessus* ou *dessous*. Nous identifions la même tendance dans l'usage de SERMO 1590-1630.

¹⁵ Voici les quatre attestations, qui n'ont pas été retenues dans nos figures, toutes repérées dans SERMO : *Seigneur, n'avons nous pas jetté hors les Diables*, François Murat 1641; *faire tirer hors ; action de mettre en avant ou tirer hors*, Jean Mestrezat 1653; *jetté hors les Diables*, Jean Daille 1667.

Nous pouvons conclure que l'interrogation des corpus permet d'affirmer que la recatégorisation des formes composées *dedans* / *dessous* / *dessus* / *dehors* en fonction de préposition s'est faite dans un espace temporel relativement bref et qu'elle est achevée probablement autour de 1630. Dans ses *Remarques*, Vaugelas fait souvent une distinction entre les usages contemporains et l'usage de la génération précédente (Ayres-Bennett 2007), et l'exemple ici décrit semble être typique de cet écart. Le texte de Leonora Christina, qui a appris le français dans une localisation périphérique par rapport au centre d'innovation parisien, témoigne que cet usage était largement diffusé et établi bien avant 1673. D'ailleurs, un examen attentif des résultats montre que la persistance de l'emploi prépositionnel de ces formes est souvent liée à des situations de figement. Le cas de SERMO 1631-1642 est d'un intérêt particulier de ce point de vue : l'influence du style biblique archaïsant particulièrement marquée chez un prédicateur explique l'usage de la construction désormais vieillie.

Pour résumer, nos recherches reflètent bien l'élimination des formes polyfonctionnelles et la séparation des usages prépositionnels et adverbiaux. Quant à l'hypothèse de périodisation avancée par Ayres-Bennett et Caron (2016), qui situent vers 1630 le point de basculement entre français préclassique et français classique, hypothèse que nous nous sommes proposé de vérifier, elle se trouve confirmée dans les grandes lignes par l'analyse de nos corpus dans ce cas précis. Il est également à noter que la distance géographique de Paris ne semble jouer aucun rôle dans ce cas.

3. Deuxième étude de cas : la montée des clitiques

3.1 Introduction

La place de certains éléments dans la phrase connaît une variation importante pendant la période explorée. Dans ce qui suit, nous nous occupons des modifications survenues concernant la position des clitiques dans des constructions verbe + infinitif et sur les effets de ces changements observés dans nos corpus pour la périodisation du français.

Le latin n'a pas à proprement parler de formes clitiques, mais des positions toniques ou non-accentuées. Selon Wackernagel (1892), la position initiale est accentuée, la deuxième position

est non-accentuée, l'élément dans cette position se rattachant de façon enclitique à l'élément initial. La distinction d'accentuation liée à la position a été réanalysée dans les langues romanes, en produisant une évolution des formes non-accentuées en cinq étapes possibles :

- 1: perte de l'accentuation
- 2: perte de l'autonomie et fixation à un hôte
- 3: fixation par rapport au verbe
- 4: enclise ou proclise par rapport au verbe
- 5: fusion.

L'histoire du français atteste le passage de l'étape 2 à l'étape 3, puis 4, avec proclise généralisée en français (clitiques devant le verbe conjugué), soit le type *je le veux voir*, alors qu'en général les autres langues romanes permettent à la fois la proclise et l'enclise, tel encore l'italien moderne : *lo voglio vedere – voglio vederlo* (cf. Bossong (1998 : 781), Maiden (2016)). Au cours de la période du français classique se manifeste une tendance à placer le clitique devant l'infinitif régi: *je le veux voir* → *je veux le voir* (cf. Haase (1930: 417-8), Galet (1971), Gougenheim (1974: 71), Fournier (1998: 80), Ayres-Bennett et Caron (2016)), mais ce n'est pourtant pas le cas avec l'ensemble des verbes régissant, puisque *faire*, *laisser* et les verbes de perception demandent jusqu'à nos jours la montée des clitiques, tel *je le vois venir*, et pas **je vois le venir*. Comment décrire et analyser cette évolution particulière au français qui a suscité beaucoup d'attention de la part des linguistes ?

Les changements survenus en français concernant la montée du clitique ont été très débattus, en particulier parce que ces changements sont contraires au chemin prévisible de la grammaticalisation (cf. Muller (1998), Crysmann (2003), Traugott et Trousdale (2014), Lehmann (2015)). En effet, il est logique de considérer les constructions avec verbes modaux et autres verbes acceptant en français préclassique et classique la montée des clitiques comme des verbes grammaticalisés en (semi-)auxiliaires (cf. le tableau de Lehmann (2015: 39) reproduit ci-dessous comme Figure 2). Ces verbes (semi-)auxiliaires se caractérisent par une structure monoclausale avec une cohésion forte entre l'auxiliaire et l'infinitif. Formellement, cette cohésion se manifeste par des contraintes de position, telles que, justement, la montée des clitiques, la position de la

négation devant l’auxiliaire et non pas devant l’infinitif, et par des règles spécifiques de reprise pronominale : *j’entends Marie chanter la chanson, je la lui entends chanter, je ne la lui entends pas chanter, j’entends cela, cela* étant la reprise de la nexie complément du verbe *entendre*, nexie composée de l’ensemble *Marie – chanter*.¹⁶ En se grammaticalisant en tant que (semi-)auxiliaires, les verbes modaux et les verbes causatifs (*faire, laisser*, etc.) suivent jusqu’à la période du français classique le chemin prévisible de la grammaticalisation (Figure 2).

[INSÉRER LA FIGURE 2 ICI]

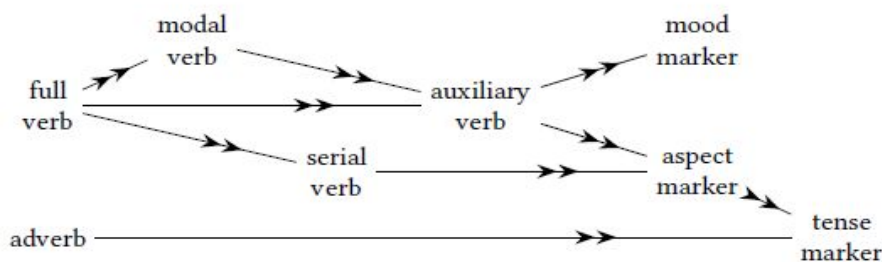


Figure 2: Les chemins de la grammaticalisation provenant de Lehmann (2015: 39)

On peut donc s’attendre à un maintien de la structure en montée des clitiques (CVInf) illustrée dans (1), suivant en cela les règles concernant la place du pronom devant l’auxiliaire des temps composés comme dans le cas de (2) :

- (1) *je le veux faire*
- (2) *je l’ai vu*

¹⁶ Une nexie est une structure dont les deux parties composantes sont dans une relation de solidarité ou d’équipollence, et pas dans une relation de subordination.

Néanmoins, cela ne s'est pas produit, puisque seule une partie relativement restreinte de l'ensemble des verbes (semi-)auxiliaires a préservé la montée des clitiques, à savoir les verbes de perception et les verbes de causation, alors que les autres, les verbes modaux et les verbes de mouvement, ont remplacé l'ordre CVInf par CVInf. Ce changement peut légitimement être analysé comme une perte de cohésion entre le verbe (semi-)auxiliaire régissant et l'infinitif, et par conséquent comme une perte du statut grammaticalisé du verbe régissant. Ainsi, le passage de l'ordre CVInf en VCIInf est un cas de *dégrammaticalisation* (Schøsler, 2007), ou bien, puisqu'il s'agit du statut d'une construction, un cas de *déconstructionnalisation*. Le fait de préserver l'ancienne structure CVInf dans un nombre restreint de contextes (*faire, laisser* et les verbes de perception), est typique pour une structure en train de disparaître qui se fait remplacer par un autre (cf. Traugott et Trousdale, 2014 : 230). Existe pourtant une autre analyse, à savoir une *regrammation* (cf. Andersen (2008), et la section 3.4.2). Dans ce qui suit, nous allons prendre en considération d'abord les réflexions des contemporains sur le changement, puis les différentes explications envisageables, avant d'explorer nos corpus.

Il est intéressant d'observer que Vaugelas (1647 : 376–377) est favorable à la construction grammaticalisée, avec la position du clitique devant l'infinitif (CVInf). D'après ses remarques, elle est beaucoup plus fréquente et a plus de grâce que celle qui va devenir la norme du français moderne:

Cette remarque est de grande estenduë; car à tous propos il s'offre occasion de dire l'un ou l'autre en d'autres exemples, que celui que je viens de proposer, comme *je ne le veux pas faire*, ou *je ne veux pas le faire*, *ils me vont blâmer*, ou *ils vont me blâmer*, & ainsi d'une infinité d'autres, où l'on employe les pronoms personnels. Il s'agit donc de sçavoir si tous deux sont bons, & cela estant, lequel est le meilleur. On respond que tous deux sont bons, mais que si celui-là doit estre appelé le meilleur, qui est le plus en usage, *je ne le veux pas faire*, sera meilleur que *je ne veux pas le faire*, parce qu'il est incomparablement plus usité. [...] il y a plus de grace, ce me semble, en cette transposition, puisque l'Usage l'autorise [...].

Il fait mention pourtant d'une exception importante: s'il y a une ambiguïté potentielle, par exemple, quand deux verbes coordonnés exigent un régime différent, l'ordre moderne est considéré plus régulier: il est donc interdit de dire : **il ne se peut taire ny parler*.

Nous avons utilisé nos corpus afin de décrire et de comprendre la dégrammaticalisation de la structure CVInf vers VCInf, à la fois contraire à l'évolution prévisible du processus de grammaticalisation, et contraire à l'usage observé et favorisé en général par Vaugelas. Notons que – malgré de nombreuses recherches sur la place des clitiques, parmi lesquelles se distingue celle de Galet (1971) – le processus de dégrammaticalisation ou de déconstructionnalisation n'a pas encore été étudié en détail en vue de décrire les deux tendances conflictuelles et de préciser la chronologie de l'évolution dans des corpus non-littéraires comme ceux du RCFC.

3.2 Exploration du corpus, remarques préliminaires

D'un point de vue technique, l'analyse de la position des clitiques dans des constructions du type *je le veux voir / je veux le voir* est facilement transformable en expressions régulières. Dans les outils d'exploration de corpus utilisant le modèle de Tree-Tagger de PRESTO, nous avons utilisé une formule combinant des recherches de mots afin de relever l'ensemble des pronoms personnels, à savoir *en, y, me, le, vous, nous, se, te, la, lui, les*, les (semi-)auxiliaires concernés, et afin d'identifier tous les verbes à l'infinitif. Un mot vide a été inséré après l'auxiliaire, pour permettre d'inclure aussi les négations *je ne veux pas le voir* ainsi que les phrases contenant un adverbe, *je veux vraiment le voir, nous ne la voulons pas confesser*.

La recherche automatique et contrôlée effectuée sur les deux corpus les plus grands (APWCF et SERMO) a permis de trouver toutes les constructions avec ou sans un seul mot intercalé, alors que la taille plus limitée des trois autres corpus (BAILLY, QUINTE-CURCE et ULFELDT) a permis le recensement de la totalité des occurrences, y compris les cas avec plus d'un élément inséré entre auxiliaire et infinitif, comme dans l'exemple suivant provenant de Soulfour 1629 : *& qui se vouloit à faux titre emparer de leur seigneurie*.

Suite à une première série d'analyses et d'améliorations des moteurs de recherche, nous avons décidé d'affiner notre requête en établissant une distinction entre plusieurs facteurs qui

n'étaient pas pris en considération au départ, en particulier le lexème régissant. Dans le but de choisir les verbes à étudier de près, nous avons consulté Galet (1971). Pour le 17^e siècle, Galet relève chez Corneille 11 verbes acceptant CVInf: *pouvoir, aller, falloir, vouloir, venir, avoir beau, daigner, devoir, oser, savoir, venir*, alors qu'au 18^e siècle, chez Voltaire, elle n'en relève que 5: *pouvoir, aller, falloir, vouloir, venir*. Inspirées par l'étude de Galet, nous avons décidé d'examiner trois types de verbes régissant, à savoir les verbes de modalité, représentés par *vouloir, pouvoir, devoir* ; les verbes de mouvement, représentés par *venir* et *aller* ; et le verbe *voir*. Ces verbes ont été choisis essentiellement pour les deux raisons suivantes : *primo* parce qu'ils représentent les trois types de structures dont les deux premiers vont subir le changement concernant la montée, contrairement au troisième type, les verbes causatifs et de perception exemplifiés par le verbe *voir* qui la préservent, et *secundo* parce qu'il était prévisible que leur fréquence est telle que nos résultats seraient fiables.

Sont écartées les constructions progressives et les citations de la Bible qui se répètent dans SERMO (elles ne sont alors comptées qu'une fois). Les exemples suivants illustrent les deux cas. Citons d'abord Mathieu chap. 3. V. 10, source des références suivantes provenant de SERMO : *Or la coignée est déjà mise à la racine des arbres : Tout arbre donc, qui ne fait pas de bon fruit s'en va être coupé & jetté au feu*. Cette citation contient un exemple de progressif, qui ne va pas être inclus dans les statistiques : *s'en va être coupé & jetté au feu*. Les sermons se réfèrent souvent à cette citation où il est question d'abattre l'arbre sans bons fruits ; nous avons jugé que leur inclusion fausserait les statistiques et nous les avons exclus. Ainsi les répétitions des mêmes schémas ne sont-ils comptés qu'une fois, par exemple :

C'étoit, dis-je, cette coignée **qui devoit les retrancher de l'Alliance de Dieu**, & leur faire perdre cet illustre privilege d'être son peuple, son héritage & son plus précieux joiau mais qu'il les menace d'un terrible jugement, **qui est de les retrancher de son alliance**, ce qui est veritablement couper l'arbre par la racine. Vengeance de Dieu, **qui devoit le retrancher de son alliance**, s'il venoit à mépriser cette grande faveur. (David Fornerod, Genève, 1688)

Nous avons envisagé la possibilité de distinguer, afin de les étudier séparément, les verbes de

mouvement à sens lexical plein et les périphrases verbales avec valeur temporelle, composées d'un verbe de mouvement en fonction d'auxiliaire suivi d'un verbe auxilié. Or, nous n'avons pas réussi à séparer les deux emplois de façon cohérente. Sont ainsi inclus dans nos statistiques des cas plutôt temporels comme :

vostre très partial Secretaire les va censurer avec une inflexible rigueur (BAILLY)

et les cas incontestables de mouvement, comme :

Le comte de Lanerich l'y va joindre avec onze mille hommes (BAILLY).

3.3 Résultats de l'exploration du corpus

3.3.1 Identification des facteurs en jeu

Nous avons exploité nos corpus afin de tester, dans la mesure où les logiciels le permettaient, l'influence de différents facteurs qui ont été proposés par les chercheurs pour expliquer la variation entre les deux structures et les éventuelles causes du changement survenu.¹⁷

Une des hypothèses proposées pour expliquer le changement de structure est celle de Rowlett (2007: 162, note 20), selon qui le processus de dégrammaticalisation serait favorisé par la présence de la négation : « in the degrammaticalised, biclausal structure [...], negation can be related to the infinitive independently of the pseudo-modal ». Cette hypothèse se heurte néanmoins à au moins deux problèmes. Tout d'abord, elle n'est pas confirmée par les faits. Une éventuelle corrélation entre l'ordre des mots VCInf et la négation a été testée en deux temps : premièrement sur le grand corpus d'APWCF, avec un résultat négatif ; deuxièmement sur les traductions de QUINTE-CURCE, avec un résultat intéressant, car dans les deux premières traductions, celles de 1598 et de 1629, un contexte négatif n'accepte que la montée puis, après 1653, c'est le contexte négatif qui est le plus favorable à l'ordre VCInf. En d'autres termes, le

¹⁷ Les analyses de SERMO sont faites avec CQPWeb, adapté à nos besoins, sur le site SERMO (<http://sermo.unine.ch/SERMO/documentation/index?part=outils>); les autres corpus utilisent des plateformes de textométrie, txm et LGerM et D* Corpora (cf. Références).

contexte négatif ne peut être la cause du changement, mais une fois le changement en cours, il pourrait favoriser le processus de changement. Le deuxième problème contre l'hypothèse citée est de nature plus fondamentale, car tout argument visant à expliquer un changement survenu à un moment précis de l'histoire doit logiquement faire appel à un ou à des facteurs nouveaux, et non à des phénomènes présents dans la langue depuis longtemps, comme la présence de la négation.¹⁸

Galet (1971) a compris la nécessité d'identifier un facteur précédant de peu le changement sous discussion. Elle soutient que le changement a été motivé par le besoin d'éviter l'ambiguïté due à la disparition du *-r* final des infinitifs. En effet, au 17^e siècle, on prononçait de la même façon non seulement *allée* et *aller* mais aussi *fini* et *finir* (Vaugelas, 1647 : 198, Galet, 1971 : 405), ce qui a pu entraîner des ambiguïtés, en particulier avec les infinitifs transitifs du type *je le veux puni(r)*, pour lesquels le pronom a pu être interprété comme sujet logique ou bien comme complément de *punir*, ce qui donne des sens très différents. L'hypothèse se heurte à un nombre de questions importantes. Tout d'abord, l'ambiguïté est-elle bien un moteur d'évolution ? La confusion entre le participe passé et l'infinitif est réelle pour les verbes transitifs en *-er*, et pour une partie des verbes en *-ir*, mais est-elle vraiment suffisante pour déclencher une réorganisation de l'ordre des mots ? La confusion est d'ailleurs inexistante pour les verbes régissants intransitifs, comme les verbes de mouvement, cf. l'exemple cité plus haut : *Le comte de Lanerich l'y va joindre avec onze mille hommes* (BAILLY) où il ne peut y avoir confusion : *le* est le complément du verbe *joindre*.¹⁹ Et enfin : l'ordre VCInf est-il vraiment plus clair que l'ordre CVInf ? Citons un exemple d'APWCF : *l'espérance de la voir continuer*. Avec l'ordre VCInf, on aurait la construction hypothétique *l'espérance de voir la continuer*, avec la même ambiguïté concernant la fonction de *la*, sujet logique ou complément de *continuer*.

¹⁸ Pourtant, un phénomène déjà existant peut gagner en importance à cause de l'apparition d'un nouveau facteur. Ceci a été examiné dans Schøsler (1984), mais une telle situation n'infirme en rien l'affirmation ci-dessus.

¹⁹ Si l'ambiguïté était un moteur de changement, on s'attendrait à une évolution différente pour les cas potentiellement ambigus (verbes régissants transitifs) et les cas non-ambigus, tels que les verbes de mouvement (cf. Wasow, Perfors et Beaver 2005). Or, nos chiffres ne permettent pas de relever une différence significative entre ces deux cas de figure, et il ne semble pas non plus que Galet (1971) en ait trouvée. Celle-ci constate, en se fondant sur l'investigation détaillée sur Corneille (ibid. : 415) que « l'évolution de la place du pronom personnel n'est pas due à la nature du verbe régent ». Par contre, au 18^e siècle, la construction ancienne (CVInf) persisterait surtout dans les cas non-ambigus (ibid. : 419).

Vaugelas, qui reconnaît l'amuïssement de *-r* final, est favorable à l'ordre CVInf, malgré la possibilité d'ambiguïté, qui n'a pas dû se présenter avec une fréquence assez grande pour donner lieu à un commentaire de sa part. D'autre part, l'hypothèse de Galet ne permet pas d'expliquer pourquoi les verbes de perception et verbes causatifs préservent l'ordre CVInf, malgré les problèmes d'ambiguïté. On comprend difficilement pourquoi l'ambiguïté du type *je le veux puni(r)* aurait pu déclencher une réorganisation importante de l'ordre des clitiques, alors que celle toute aussi réelle du type *je le vois puni(r)* n'a pas eu le même effet. Suite à ces réflexions, nous allons explorer notre corpus dans le but de nous éclairer sur le point suivant : si l'ambiguïté est le moteur du changement, il faut prévoir une variation entre CVInf et VCInf dans le cas aussi du troisième type de verbes, les verbes de perception²⁰ et les verbes causatifs. Reste en outre à expliquer pourquoi CVInf aurait été rétabli ou préservé avec ce type de verbes. S'il n'y a pas de variation entre les deux structures pour les verbes de perception, l'hypothèse de Galet sur l'importance de la possibilité de confusion pour l'abandon de la montée pourra difficilement être retenue.

3.3.2 Statistiques

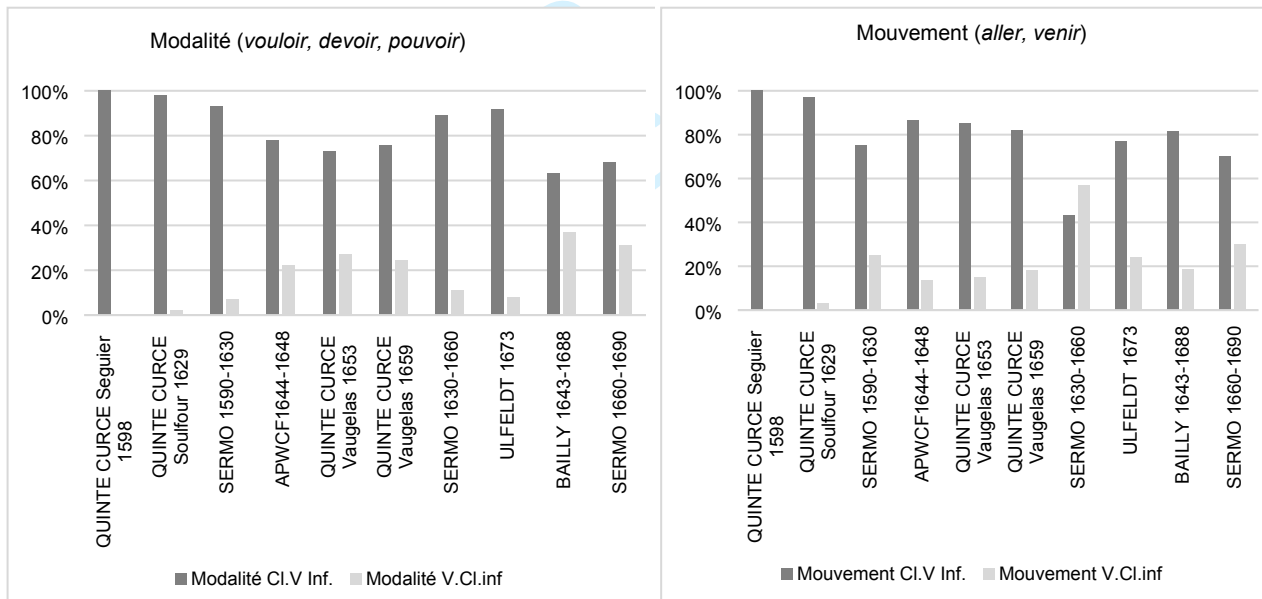
Voici les résultats de l'interrogation sur les sous-corpus pour ce changement (Tableau 3 et Figure 3):

[INSÉRER ICI LE TABLEAU 3 et LA FIGURE 3]

²⁰ Nous avons choisi d'explorer le verbe *voir* en tant que représentant des verbes de perception, étant donné que c'est le verbe de ce type le plus fréquent.

Verbes		QUINTE CURCE Seguier 1598	QUINTE CURCE Souffour 1629	SERMO 1590-1630	APWCF 1644-1648	QUINTE CURCE Vaugelas 1653	QUINTE CURCE Vaugelas 1659	SERMO 1630-1660	ULFELDT 1673	BAILLY 1643-1688	SERMO 1660-1690
Modalité (Vouloir, pouvoir, devoir)	Clitique + Verbe + Infinitif	82	48	132	5918	54	50	106	67	422	176
	Verbe + Clitique + Infinitif	0	1	10	1690	20	16	13	6	246	82
Mouvement (aller / venir)	Clitique + Verbe + Infinitif	7	35	3	216	23	18	3	10	166	7
	Verbe + Clitique + Infinitif	0	1	1	34	4	4	4	3	38	2
Sensation (voir)	Clitique + Verbe + Infinitif	0	0	1	83	0	0	1	0	15	2
	Verbe + Clitique + Infinitif	0	0	0	3	0	0	1	0	0	0
Total	Clitique + Verbe + Infinitif	89	83	136	6217	77	68	110	77	603	185
	Verbe + Clitique + Infinitif	0	2	11	1727	24	20	18	9	284	84

Tableau 3: Constructions avec clitiques dans les corpus du RCFC pour les verbes de modalité (*vouloir, pouvoir, devoir*), de mouvement (*venir, aller*), du verbe *voir* et total



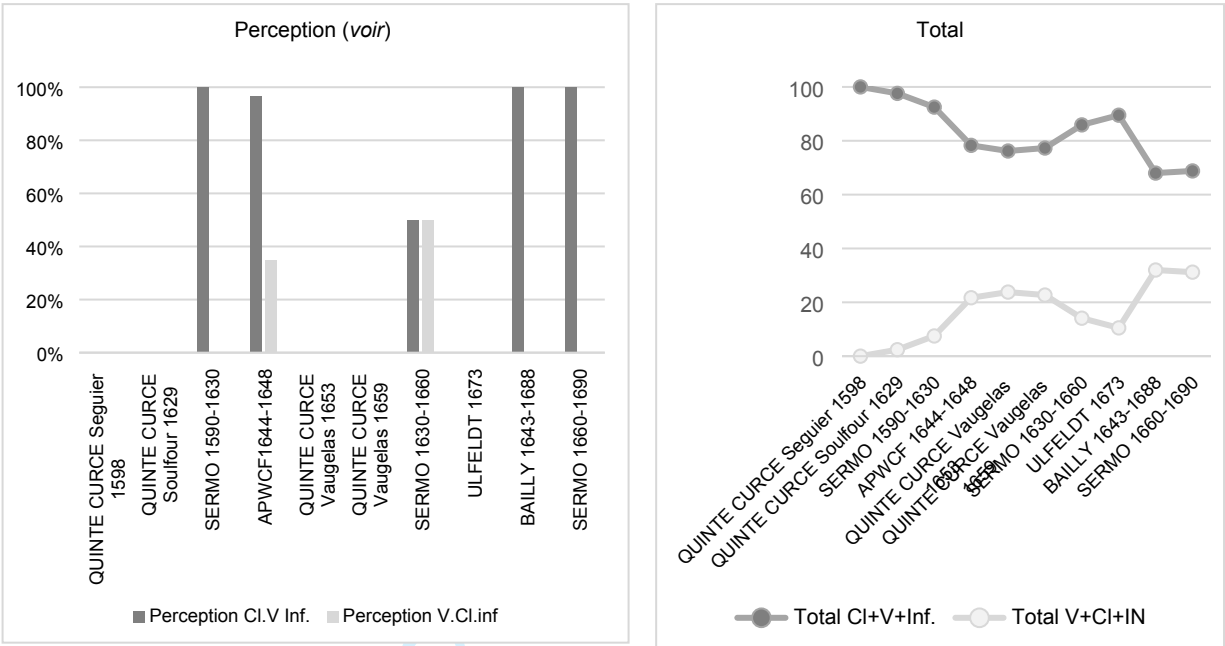


Figure 3: Constructions avec clitiques dans les corpus du RCFC pour les verbes de modalité (*vouloir*, *pouvoir*, *devoir*), de mouvement (*venir*, *aller*), du verbe *voir* et total

3.4 Analyses

3.4.1 Tendances générales

Quelques tendances se dégagent nettement de nos recherches. En premier lieu, comme on pouvait s'y attendre, les corpus montrent globalement une progression diachronique en faveur de l'ordre VCInf, à l'exception du verbe *voir*. Le type de verbe est par conséquent le facteur déterminant pour la distribution entre CVInf et VCInf: les verbes de modalité et les verbes de mouvement se comportent de façon assez comparable, avec quelques écarts entre les sous-corpus. Par contre, le verbe *voir*, qui représente un verbe de perception, ne participe aucunement au changement en cours – nous pouvons l'affirmer malgré la faible présence de *voir* avec pronom et infinitif. Le seul exemple relevé de VCInf du corpus SERMO avec *voir* est fort complexe, avec un complément d'objet direct relatif antéposé (*laquelle*), complément du verbe réfléchi *se faire*. La constellation **laquelle nous se voyons faire* n'est pas possible : *comme de l'illumination de l'air, laquelle nous voyons se faire en un instant*. Les quelques cas trouvés dans APWCF s'expliquent également par la non-compatibilité des pronoms,²¹ tel que celui-ci: *et selon le jour qu'il y verra luy faire mesme instance de retirer ses troupes de l'armée de l'Empereur* (1647). Nous commenterons plus bas la différence entre les trois types de verbes ci-dessous (3.4.2).

En deuxième lieu, parmi les verbes qui commencent à adopter l'ordre VCInf, le facteur temps est déterminant. Nos recherches montrent une modification progressive allant d'une utilisation quasi-exclusive de la structure CVInf avec les verbes de modalité et les verbes de mouvement jusque vers 1630, suivi par un recul lent en faveur de la structure VCInf, encore minoritaire dans l'ensemble de nos corpus, mais qui apparaît nettement vers 1660 (cf. Figure 3). Ce changement se manifeste surtout dans les corpus longitudinaux (QUINTE-CURCE, SERMO), dont les données anciennes (intégrées dans Figure 3) sont les suivantes: dans Séguier 1598 l'ordre CVInf se rencontre dans 100% des cas; dans Souffour 1629 CVInf se rencontre dans 97,6%,

²¹ Certains pronoms, y compris les clitiques *me*, *te* *se*, *nous* et *vous* ne peuvent pas se trouver juxtaposés deux à deux, et il y a des contraintes avec *y* et *en* en seconde position, donc **nous vous voir faire*, **voir nous vous faire*, **voir lui y faire* ou **lui y voit faire* sont exclus.

et dans les sous-corpus SERMO entre 1590 et 1630 dans 92,5%. La chronologie différente des corpus explique quelquefois les différences de pourcentage. Si, par exemple BAILLY apparaît à première vue plus conservateur que les autres corpus, en regardant de plus près on observe que les exemples de VCInf sont de plus en plus nombreux au cours des années et surtout à partir des années 1660 (à titre d'exemple pour le verbe *vouloir* on enregistre 28,7 % de VCInf jusqu'en 1659 et 36,3 % entre 1660 et 1673). D'autre part ULFELDT est plus conservateur que les autres textes.

Pour ce qui est de la progression de l'ordre VCInf avec les verbes modaux ou les verbes de mouvement, chaque corpus semble se comporter différemment. Dans certains corpus, les verbes de modalité semblent plus avancés (APWCF, Vaugelas 1653, 1659, BAILLY), alors que dans d'autres (notamment ULFELDT) les verbes de modalité sont en recul par rapport à l'évolution. Nous reviendrons à cette question dans la section 3.4.3 ci-dessous.

La figure 3 nous permet de classer nos corpus selon qu'ils suivent l'une ou l'autre des deux tendances contrastées: la progression prévisible de la grammaticalisation (CVInf) d'une part, ou celle du nouveau standard (VCInf) d'autre part, avec dégrammaticalisation de l'ancienne structure (CVInf).²² Il est intéressant d'observer que Vaugelas, à qui on attribue souvent une influence décisive sur la norme moderne, donne explicitement sa préférence à la structure CVInf, et persiste à s'en servir dans ses deux traductions.

3.4.2 Commentaires sur le type de verbe

Il est opportun d'analyser de plus près le comportement du verbe *voir* comparativement aux deux types de verbe qui vont subir le processus de déconstructionnalisation de CVInf à VCInf. Comment expliquer le comportement différent de *voir*? Rappelons qu'il y a une différence importante concernant les verbes de perception par rapport aux verbes de modalité et de mouvement, car les verbes causatifs et les verbes de perception partagent la propriété d'ajouter un actant à la valence

²² Les linguistes qui travaillent sur la grammaticalisation sont en désaccord sur l'existence de processus de dégrammaticalisation ou *degrammation*. Cf. entre autres deux points de vue contrastés dans Andersen (2006, 2008) et Prévost (2003). Comme il s'agit ici d'une structure grammaticalisée (CVInf) qui perd son statut grammaticalisé, nous utilisons ici le terme de dégrammaticalisation. Toutefois, dans la mesure où il s'agit de la perte de la valeur grammaticale d'une construction dans le sens de la grammaire des constructions (Construction Grammar), il serait plus juste d'utiliser le terme de *déconstructionnalisation*.

de l'infinitif (cf. Tesnière, 1976: chapitre 108), ce qui implique que le sujet du verbe régissant n'est pas identique à celui (logique) de l'infinitif. Alors que pour les deux autres types, il y a identité de sujet du verbe régissant et de celui, logique, de l'infinitif. Cette différence ne peut évidemment pas être un moteur de changement, puisqu'elle n'est pas nouvellement introduite, mais elle pourrait contribuer à expliquer pourquoi ce type, essentiellement différent des autres, malgré la ressemblance superficielle de la position du clitique par rapport au verbe, présente la particularité de ne pas suivre le processus de déconstructionnalisation de CVInf à VCInf survenue avec les deux autres types. Cette différence pourrait être interprétée de la façon suivante: au lieu d'être perçu comme marqueur de (semi-)auxiliaire, l'ordre CVInf est regrammaticalisé (selon la terminologie d'Andersen, 2008) comme marqueur d'augmentation de valence. En outre, la persistance de CVInf avec les verbes de perception implique que l'hypothèse de Galet sur l'influence de l'amuïssement de *-r* final pour la disparition de la construction CVInf ne pourra pas être retenue.²³

3.4.3 Commentaires sur la différence de comportement des sous-corpus

Si nous considérons les traductions de Vaugelas comme représentant la norme émergente (VCInf ca. 23%), il est frappant de noter que les chiffres pour SERMO (ca. 20%) et APWCF (ca. 22%) concordent dans les grandes lignes. Autrement dit, dans ces deux cas, le changement de genre ne semble pas modifier la chronologie du changement, contre nos attentes.

Comme certains textes sont rédigés par des étrangers, peut-on déceler une influence diatopique (centrifuge) sur le choix de la structure? Cette question est problématique, car la plupart de nos auteurs ont été en relation permanente avec le centre de diffusion de la norme qui est constitué de Paris et de la Cour. Il ne faut pas oublier que Vaugelas lui-même était savoyard de naissance, même s'il a passé la plupart de sa vie adulte à la cour. Bailly était également savoyard, puis installé à Paris. Le cas de ULFELDT est toutefois intéressant de ce point de vue. Comme

²³ Nous n'avons pas été en mesure d'explorer d'autres facteurs liés aux verbes individuels, régissants ou régis, déjà largement étudié par Galet, car notre but est d'abord de montrer l'avantage d'inclure dans l'étude de la période des corpus non encore étudiés, et ensuite de tester l'hypothèse concernant la chronologie des changements.

nous l'avons vu dans la section 2, la princesse a intégré dans sa langue la distinction préposition / adverbe, changement qui s'est produit vers 1630. Après 1647, elle n'a plus fréquenté la cour royale, et elle a été coupée de son cercle francophone depuis son incarcération en 1663, date vers laquelle le changement concernant la montée des clitiques se manifeste plus nettement. Par conséquent il est probable que son éloignement de la cour explique au moins en partie la nature conservatrice de son usage en ce domaine.

4. Conclusion générale

Pour cette enquête nous avons choisi deux études de cas qui représentent des exemples importants relevant de la restructuration de la grammaire ; abondamment discutés par les remarqueurs, ils sont saillants à l'époque et constituent ainsi des cas types pour diagnostiquer le changement linguistique dans les premières décennies du 17^e siècle. Le premier concerne la recatégorisation ; le second peut être interprété aussi comme une tendance vers la création d'une catégorie plus large d'auxiliaires avec, comme trait spécifique, la montée des clitiques, tendance qui subit un changement important vers 1660. La préservation de la montée avec un petit nombre de verbes peut être interprétée comme une sorte de fossilisation de la construction CVInf ou bien comme une regrammation.

Du point de vue de l'évolution, il faut noter que les deux changements que nous avons examinés se sont produits à des moments différents, ou plutôt qu'ils se sont manifestés dans deux étapes différentes de l'évolution, le premier représentant la dernière étape du changement, et le second les premières étapes. La distinction entre préposition et adverbe (*dans / dedans*, etc.) s'achève autour de 1630, tandis que la montée des clitiques commence seulement à s'accélérer à partir de cette même date. Ils nous donnent ainsi deux perspectives différentes sur la question de savoir s'il y a une frontière entre deux chronoclectes autour de 1630.

En considérant la diachronie des changements examinés et ainsi la question de la périodisation, il est peut-être surprenant que la prise en compte de textes de types variés, et pas seulement de type littéraire, ne semble pas avoir beaucoup modifié la chronologie de l'évolution des changements déjà établie en faisant l'analyse de textes littéraires. Autrement dit, contre nos

attentes, les textes non-littéraires ne sont pas en avance par rapport aux textes littéraires. En ce qui concerne la différenciation des prépositions et des adverbes, il apparaît que la spécificité des typologies textuelles n'a pas joué un grand rôle à l'exception près de SERMO 1631-1642 où le style de deux prédicateurs, influencé par le style biblique, explique la tendance conservatrice de ce corpus. Bien qu'il y ait variation selon la préposition, après Soulfour 1629, le niveau d'attestation de l'ancien usage est toujours inférieur à 5% dans tous les autres corpus étudiés. Ce changement, terminé vers 1630, est confirmé même dans le texte de Leonora Christina, qui l'a intégré dans son français, appris dans une localisation périphérique par rapport au centre d'innovation.

Dans le deuxième cas, celui de la montée des clitiques, nos résultats s'alignent aussi dans les grandes lignes avec ceux de Galet, qui identifie la date de 1660 comme le point de basculement vers l'ordre moderne, et souligne, par exemple, les corrections faites par Pierre Corneille dans son édition de 1662. Il est notable que le rythme du changement dans les lettres diplomatiques (APWCF) et dans les sermons (SERMO) est assez proche de celui de Vaugelas (1653, 1659), dont l'usage représente la norme émergente. Si BAILLY semble plus progressiste, en utilisant la construction moderne avec une fréquence un peu plus élevée, nous avons vu que cette différence n'est qu'apparente et s'explique par la chronologie longue de ce corpus (1643-1688). ULFELDT, d'autre part, est conservateur, malgré le fait qu'il s'agit d'un « ego-document », source souvent privilégiée par les études de sociolinguistique historique. Ici la question se complique pourtant par le fait que la Princesse est également locutrice non-native, et qu'elle avait moins de contact avec la cour et les sources du bon usage après son incarcération en 1663. Parmi les textes rédigés loin du centre (ULFELDT, BAILLY, SERMO), seul ULFELDT est en retard par rapport aux textes rédigés près du centre (QUINTE-CURCE, APWCF).

Bref, l'élargissement des types de textes étudié semble confirmer dans les grandes lignes l'hypothèse selon laquelle il y aurait une « rupture » autour de 1630, comme cela a été proposé par Ayres-Bennett et Caron (2016). Il reste à déterminer s'il existe un autre faisceau d'isoglosses autour de 1660 dont la montée des clitiques fait partie, ou si c'est un changement plutôt isolé. Pour préciser encore la datation de changements linguistiques importants se produisant au 17^e siècle,

il faudra multiplier les études, en considérant d'autres faits linguistiques et en nous appuyant sur d'autres types de textes.

For Peer Review

Références

- Amatuzzi, A., Ayres-Bennett, W., Gerstenberg, A., Schøsler, L. et Skupien-Dekens, C. (2019). Améliorer et appliquer les outils numériques. Ressources et approches pour l'étude du changement linguistique en français préclassique et classique. In : D. Capin, J. Glikman, V. Obry et T. Revol (dirs), *Le français en diachronie. Moyen français – Segmentation des énoncés – Linguistique textuelle*, Strasbourg : Société de Linguistique Romane / Éditions de linguistique et de philologie, pp. 337-364.
- Ayres-Bennett, W. (2007). Vaugelas's Literary Sources: *Anciens, Modernes and Demi-Modernes*. *Zeitschrift für französische Sprache und Literatur*, 117(3): 225-238.
- Ayres-Bennett, W. et Caron, P. (1996). *Les Remarques de l'Académie Française sur le Quinte-Curce de Vaugelas 1719-1720*. Paris: Presses de l'Ecole normale supérieure.
- Ayres-Bennett, W. et Caron, P. (2016). Periodization, Translation, Prescription, and the Emergence of Classical French. *Transactions of the Philological Society*, 114(1): 339–390.
- Ayres-Bennett, W. et Seijido, M. (2011). *Remarques et observations sur la langue française : histoire et évolution d'un genre*. Paris: Éditions Classiques Garnier.
- APWCF-DIACOLLO = Gerstenberg, A. et al. (2017). *Acta Pacis Westphalicae, Plateforme Diacollo*. FU Berlin, BBAW:<http://kaskade.dwds.de/dstar/apwcf/diacollo>.
- Gerstenberg, A. (2018). À l'intérieur du genre, les sources françaises des *Acta Pacis Westphalicae*. Approches linguistiques. In: W. Ayres-Bennett, A. Carlier, J. Glikman, Th. M. Rainsford, G. Siouffi et C. Skupien Dekens (dirs), *Nouvelles voies d'accès au changement linguistique*. Paris: Classiques Garnier, pp. 107–121.
- Schøsler, L. (1984). *La déclinaison bicasuelle de l'ancien français, son rôle dans la syntaxe de la phrase, les causes de sa disparition*. Odense : Etudes romanes de l'Université d'Odense.
- Schøsler, L. (2007). Grammaticalisation et dégrammaticalisation: Étude des constructions progressives en français du type Pierre 'va / vient / est chantant'. In : E. Labeau, C. Vettters

et P. Caudal (dirs), *Sémantique et diachronie du système verbal français* (Cahiers Chronos 16). Amsterdam / New York: Rodopi, pp. 91–119.

Skupien-Dekens, C. (2017). Du Corbeau enrôlé au Patois de Canaan, l'influence des traductions bibliques sur le sociolecte protestant. In : O. Christin et Y. Krumenacker (dirs), *Les Protestants à l'époque moderne, une approche anthropologique*. Rennes: Presses Universitaires de Rennes., pp. 397-416.

Skupien-Dekens, C. (2018). La langue de Canaan » à l'épreuve des sermons (1600-1750). L'exemple des psaumes. *Revue Bossuet*, 9, 35-58.

Andersen, H. (2006) : Grammatication, regrammation, and degrammation. Tense loss in Russian, *Diachronica* 23: 2 : 231-258

Andersen, H. (2008). Grammaticalization in a speaker-oriented theory of change. In: T. Eythórsson (dir.), *Grammatical change and linguistic theory: The Rosendal Papers*, 113. Amsterdam: John Benjamins, pp. 11-44.

Blake, N. F. (1994). Premisses and Periods in a History of English. In: F. Fernández, M. Fuster et J. J. Calvo (dir.), *English Historical Linguistics 1992: Papers from the 7th International Conference on English Historical Linguistics, Valencia, 22–26 September 1992*, Amsterdam / Philadelphia : John Benjamins, pp. 37-46.

BMF = ENS de Lyon. 2016. *Base de Français Médiéval*. Lyon: <http://bfm.ens-lyon.fr/>.

Bossong, G. (1998). Eléments d'une typologie actancielle des langues romanes. In : J. Feuillet (dir.), *Actance et Valence dans les Langues de l'Europe*. Berlin / New York: Mouton de Gruyter, pp. 769-787.

Caron, P. (2002).. Vers la notion de chronolecte ? Quelques jalons à propos du français préclassique. In : R. Sampson et W. Ayres-Bennett (dirs), *Interpreting the History of French: A Festschrift for Peter Rickard on the Occasion of his Eightieth Birthday*. Amsterdam / New York: Rodopi, pp. 329-352.

Combettes, B. et Marchello-Nizia, C. (2010). La périodisation en linguistique historique: le cas du français classique. In : B. Combettes et al. (éds), *Le changement en français. Etudes de*

linguistique diachronique. Bern / Berlin / Bruxelles / Frankfurt am Main / New York / Oxford / Wien: Lang, pp. 129–142.

Crysmann, B. (2003). Clitic Climbing Revisited, In: J.-B. Kim et S. Wechsler (dirs), *The Proceedings of the 9th International Conference on HPSG*. Stanford : CSLI Publications, pp. 67-89.

Dstar = Jurish, B. [2019], D* OpenSearch API version 0.46, Berlin-Brandenburgische Akademie der Wissenschaften, <http://kaskade.dwds.de/dstar>.

De Mulder, W. (2003). La préposition au dessus de : un cas de grammaticalisation. *Verbum*, 25: 293-307.

Diedihou, F. (2011). La grammaticalisation des prépositions : le cas de *au-dessus*. *Sud Langues*, 15: 71-85.

Fagard, B. et Sarda, L. (2009). Étude diachronique de la préposition *dans* : des emplois spatiaux-temporels aux emplois abstraits ? In : J. François (dir.), *Autour de la préposition*. Caen : Presses Universitaires de Caen, pp. 225-236.

Fagard, B. et De Mulder, W. (2007). La formation des prépositions complexes : grammaticalisation ou lexicalisation ? *Langue française*, 156: 9-29.

Fagard, B. et Combettes, B. (2013). De *en* à *dans*, un simple remplacement ? Une étude diachronique. *Langue française*, 178(2): 93-115.

Fournier, N. (1998). *Grammaire du français classique*. Paris: Belin.

Galet, Y. (1971). *L'évolution de l'ordre des mots dans la phrase française*. Paris: PUF.

Gougenheim, G. (1974). *Grammaire de la langue française du seizième siècle*, nouvelle éd. refondue. Paris: A. et J. Picard.

Haase, A. (1930). *Syntaxe française du XVII^e siècle*: Nouvelle édition traduite et remaniée par M. Obert. Avec l'autorisation de l'auteur. Paris: Delagrave.

Heiden, S., Magué, J.-P. et Pincemin, B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. *Journées internationales d'Analyse Statistique des données Textuelles 7*.

- Koch, P. et Oesterreicher, W. (1985). Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36 / 85, 15-43.
- Koch, P. et Oesterreicher, W. (2001). Langage parlé et langage écrit. In : G. Holtus, M. Metzeltin, M. et C. Schmitt (dirs), *Lexikon der Romanistischen Linguistik* I / 2. Tübingen, Niemeyer, pp. 584-627.
- Leclercq, O. (2018). Changement linguistique et diffusion de la norme aux siècles classiques : le cas de *dedans*, *dessous*, *dessus* et *dehors* prépositions. In : B. Colombat, B. Combettes, V. Raby et G. Siouffi (dirs), *Histoire des langues et histoire des représentations linguistiques*. Paris : H. Champion, pp. 481-500.
- Lehmann, C. (2015). *Thoughts on Grammaticalization* (Classics in linguistics 1), 3ème édn. Berlin: Language Science Press.
- LGeRM = [2017]. *Lemmes Graphies et Règles Morphologiques*. ATILF - CNRS et Université de Lorraine: <http://www.atilf.fr/LGeRM/>.
- Maiden, M. (2016). *Linguistic history of Italian*. London: Routledge.
- Maingueneau, D. (2009). Le sermon : contraintes génériques et positionnement. *Langage et société*, 130(4) : 37-59.
- Marchello-Nizia, C. (1995). *L'évolution du français. Ordre des mots, démonstratifs, accent tonique*. Paris: Armand Colin.
- Marchello-Nizia, C. (1999) : *Le français en diachronie : douze siècles d'évolution*, Gap : éditions Ophrys.
- Muller, C. (1998). On the Position of Romance Clitics. *Hermes, Journal of Linguistics*, 20: 115-138.
- Nevalainen, T. et Raumolin-Brunberg, H. (2012). Historical sociolinguistics: Origins, motivations, and paradigms. In: J. M. Hernández-Campoy et J. C. Conde-Silvestre (dirs) *The Handbook of Historical Sociolinguistics*. Chichester, West Sussex, UK / Malden, MA : Wiley-Blackwell, pp. 22-40.

PRESTO = *Projet ANR/DFG : l'évolution du système prépositionnel du français*. ENS de Lyon:

<http://presto.ens-lyon.fr>, 2015..

Prévost, S. (2003). La grammaticalisation : unidirectionnalité et statut, *Le français Moderne* LXXI (2). 144-166.

RCFC = Réseau Corpus Français Classique et Pré-Classique, Cambridge / Copenhague / Neuchâtel / Potsdam / Torino, 2014–2019.

<https://diachronie.org/category/ressources/corpus/>.

Rowlett, P. (2007). *The Syntax of French*. Cambridge / New York: Cambridge University Press.

SERMO = Corpus SERMO Université de Neuchâtel: <http://sermo.unine.ch/>.

TreeTagger = Schmid, H. (2008). TreeTagger – a Language Independent Part-of-speech Tagger.

<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger>

<http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger>.

Traugott, E. C. et Trousdale, G. (2013). *Constructionalization and Constructional Changes*. Oxford: Oxford Scholarship Online.

Tremblay, M., Dupuis, F. et Dufresne, M. (2003). Les prépositions dans l'histoire du français: Transitivité, grammaticalisation et lexicalisation. *Verbum*, 25(4) : 549–562.

Vaugelas, C F. de (1647). *Remarques sur la langue française*. Paris: Vve J. Camusat et P. Le Petit.

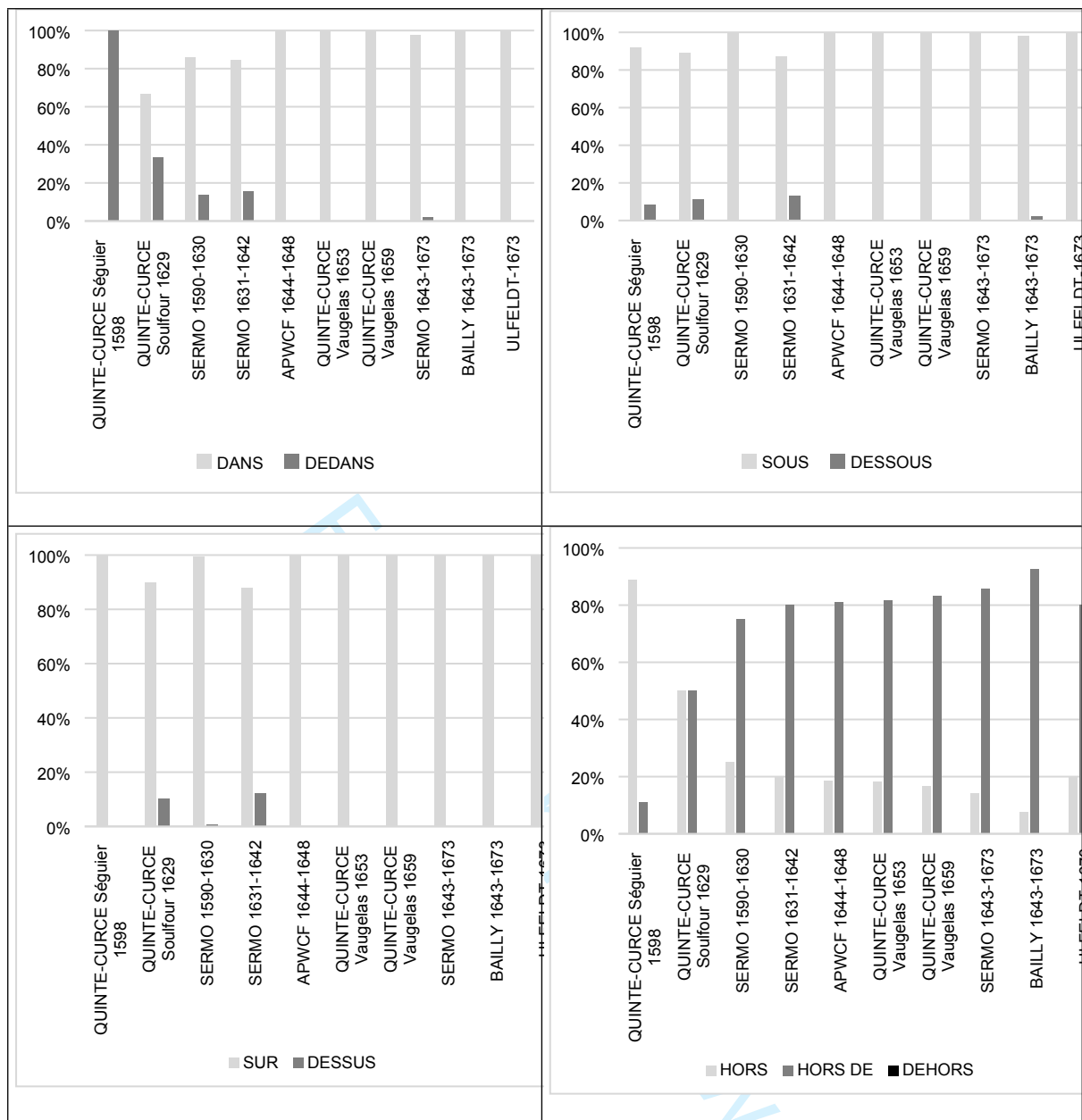
Vigier, D. (2015). Les prépositions *en*, *dans* et *dedans* au XVI^e siècle. Approche statistique et combinatoire. *Le Français Moderne*, 2 : 228–247.

Vigier, D. (2017). La préposition *dans* au XVI^e siècle. Apports d'une linguistique instrumentée. *Langages*, 206(2) *Du quantitatif au qualitatif en diachronie : prépositions françaises* : 105-122.

Wackernagel, J. (1892). Über ein Gesetz der indogermanischen Wortstellung, *Indogermanische Forschungen* 1. 333–436.

Wasow, T., Perfors, A. et Beaver, D. (2005). The Puzzle of Ambiguity. In: O. Orgun et P. Sells (dirs) *Morphology and the Web of Grammar: Essays in Memory of Steven G. Lapointe*. Stanford: CSLI Publications, pp. 265-282.

For Peer Review

Figure 1 : Formes en fonction de préposition¹

¹ On trouve une seule occurrence de *dehors* comme préposition, dans le corpus APWCF 1644-1648. La fréquence est trop faible pour apparaître dans la figure.

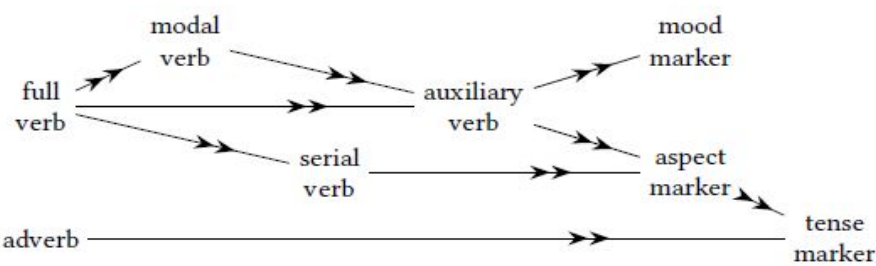


Figure 2: Les chemins de la grammaticalisation provenant de Lehmann (2015: 39)

For Peer Review



Figure 3: Constructions avec clitiques dans les corpus du RCFC pour les verbes de modalité (*vouloir, pouvoir, devoir*), de mouvement (*venir, aller*), du verbe *voir* et total.

	Communication		Actionnel et situationnel		Spatio-temporel		Coopération communicative		Communication	
	privée	publique	ancrage	détachement	co-présence	séparation	intense	minime	spontanée	préparée
SERMO		+	+		+	+		+		+
QUINTE-CURCE		+		+		+		+		+
BAILLY	+	+	+			+	+	+	+	
APWCF	+	+	+	+		+		+	+	+
ULFELDT	+		+			+	+	+	+	

Tableau 1 : Sélection des paramètres de la communication (Koch & Oesterreicher, 2001 : 586)

For Peer Review

	QUINTE- CURCE Ségurier 1598	QUINTE- CURCE Souffour 1629	SERMO 1590-1630	SERMO 1631-1642	APWCF 1644-1648	QUINTE- CURCE Vaugelas 1653	SERMO 1643-1673	BALLY 1643-1673	QUINTE- CURCE Vaugelas 1659	ULFELDT 1673
<i>Dans + N ou Pro</i>	0	100	31	60	11677	176	527	1705	174	75
<i>Dedans+ N ou Pro</i>	64	50	5	11	6	0	12	0	0	0
<i>Sous+ N ou Pro</i>	11	8	33	20	216	30	82	51	21	2
<i>Dessous+N ou Pro</i>	1	1	0	3	0	0	0	1	0	0
<i>Sur+N ou Pro</i>	109	149	140	43	8421	162	336	1066	143	71
<i>Dessus+ N ou Pro</i>	0	17	1	6	9	0	0	0	0	0
<i>Hors+ N ou Pro</i>	8	6	3	1	70	2	2	4	2	2
<i>Hors de + N ou Pro</i>	1	6	9	4	303	9	12	49	10	8
<i>Dehors + N ou Pro</i>	0	0	0	0	1	0	0	0	0	0

Tableau 2 : Formes en fonction de préposition

Verbes		QUINTE CURCE Seguler 1598	QUINTE CURCE Souffour 1629	SERMO 1590-1630	APWCF 1644-1648	QUINTE CURCE Vaugelas 1653	QUINTE CURCE Vaugelas 1659	SERMO 1630-1660	ULFELDT 1673	BAILLY 1643-1688	SERMO 1660-1690
Modalité (Vouloir, pouvoir, devoir)	Clitique + Verbe + Infinitif	82	48	132	5918	54	50	106	67	422	176
	Verbe + Clitique + Infinitif	0	1	10	1690	20	16	13	6	246	82
Mouvement (aller / venir)	Clitique + Verbe + Infinitif	7	35	3	216	23	18	3	10	166	7
	Verbe + Clitique + Infinitif	0	1	1	34	4	4	4	3	38	2
Sensation (voir)	Clitique + Verbe + Infinitif	0	0	1	83	0	0	1	0	15	2
	Verbe + Clitique + Infinitif	0	0	0	3	0	0	1	0	0	0
Total	Clitique + Verbe + Infinitif	89	83	136	6217	77	68	110	77	603	185
	Verbe + Clitique + Infinitif	0	2	11	1727	24	20	18	9	284	84

Tableau 3: Constructions avec clitiques dans les corpus du RCFC pour les verbes de modalité (*vouloir, pouvoir, devoir*), de mouvement (*venir, aller*), du verbe *voir* et total